Response to Buchwalter et al.: Further Considerations for Modernising Water Quality Criteria in the USA and Elsewhere

Warne, M, Van Dam, R, Batley, G & Stauber, J

Original citation & hyperlink:
Warne, M, Van Dam, R, Batley, G & Stauber, J 2017, 'Response to Buchwalter et al.: Further Considerations for Modernising Water Quality Criteria in the USA and Elsewhere.' *Environmental Toxicology and Chemistry*, Vol 36, Issue 6, pp. 1422-1424
http://dx.doi.org/10.1002/etc.3812

ISSN 0730-7268
ESSN 1552-8618

Publisher: Wiley

Response to Buchwalter et al. (2017): Further Considerations for Modernising Water Quality Criteria in the USA and Elsewhere

Warne MStJ [†, ‡, §, ||], Van Dam R[#], Batley GE[††], Stauber JL[††]

[†]. Centre for Agroecology, Water and Resilience, Coventry University, Coventry, West Midlands, United Kingdom. [‡]. Queensland Department 12of Science, Information Technology and Innovation, Brisbane, Queensland, Australia. [§]. Australian Rivers Institute, Griffith University, Brisbane, Queensland, Australia. [||]. National Research Centre for Environmental Toxicology, Brisbane, Queensland, Australia. [#]. Environmental Research Institute of the Supervising Scientist, Department of the Environment, Darwin, Northern Territory, Australia. [††]. Commonwealth Scientific and Industrial Research Organisation (CSIRO), Land and Water, Lucas Heights, NSW, Australia.

Keywords: water quality criteria, water quality guidelines, contaminants

Running title: Response to Buchwalter et al. (2017)

The Buchwalter et al. ET&C Focus article [1] is a very timely piece given the review of the method used by the USEPA to derive water quality criteria (WQC; [2]), and also because of the ongoing push for increased international harmonization of derivation methods for WQC (and their analogues in other jurisdictions) (e.g. [3, 4, 5]). Buchwalter et al. [1] provide a cogent argument not only for the expansion of the definition of data that are acceptable to derive WQC, but also for a more holistic paradigm to deriving WQC.

While there are distinct differences between the definitions of WQC, guideline values (GVs), and environmental quality standards (EQSs), for simplicity we will refer to them collectively as benchmarks, except when citing specific examples.

Buchwalter et al. [1] propose a new paradigm “whereby all relevant knowledge (from both lab and field) is brought to bear to reduce uncertainty as to whether criteria meet the stated goal of protecting aquatic ecosystems.” We concur that a weight of evidence approach to the derivation of benchmarks is the ideal and most rigorous approach, however, we feel that there are some additional considerations to add to the debate. These are discussed below.

The central tenets of our response are: (i) while the focus of Buchwalter et al. [1] is on the US regulatory process, a number of their suggestions have already been adopted in other jurisdictions; and (ii) while we support most of their recommendations, we believe that it is crucial to achieve the right balance between a goal of ever-increasingly accurate benchmarks for a very limited number of chemicals (e.g., the USA has WQC for approximately 40 chemicals) and having less accurate default benchmarks for far more chemicals (e.g., Australia and New Zealand have GVs for over 300 chemicals). Having benchmarks for a large number of relevant chemicals together with a robust system that incorporates the types of recommendations made by Buchwalter et al. [1], is needed to accurately assess water quality in a site-specific or regional context.

Buchwalter et al. [1] recommend that data for more types of organisms be used, but they do not explain how this should be achieved. Rather, they state that the examples they provided “should compel USEPA to broaden the coverage of various faunal groups in toxicity assessments used in WQC development”. We also support the call for an expansion of the types of organisms used to derive benchmarks, but do not support the expansion of a prescribed list of organism types, for two key reasons: (i) species’ sensitivities will differ between chemicals due to many factors including species’ life histories and traits as well as physico-chemical properties and mode of action of chemicals, making it almost impossible to prescribe an adequate coverage of faunal groups, and (ii) species sensitivity distribution (SSD) and genus sensitivity distribution (GSD) methods are statistical methods and, as such, assume that the organisms used are a random selection of all species in the environment being protected [e.g., 6, 7, 8]. While it is clear that the species used in ecotoxicity tests are not a random subset of species (but rather have been selected for pragmatic reasons including that they can be maintained and bred in the laboratory, have relatively short life cycles, etc.) prescribing the types of organisms that must be represented further invalidates this assumption. The same effect of increasing the number of organism types can be achieved by setting higher minimum data requirements with some degree of prescription (e.g., the EU minimum data requirements are for 10 species that belong to eight taxonomic groups and a list of taxa that "would normally need to be represented" [9] or simply specifying the minimum number of species and taxonomic groups with no specific taxa required [e.g., 10, 11]. Moreover, it would be beneficial for those generating toxicity data, to select species based on a conceptual understanding of the chemical's properties, mode of action and likely organism susceptibilities, rather than just the standard toxicity testing species used by commercial and many research laboratories. While doing this, the aim of using data from as many species as possible from the most diverse range of organism types should still be the goal.

In proposing their paradigm, Buchwalter et al. [1] espouse the benefits of using field and/or mesocosm data. Both Australia and New Zealand and the EU have long permitted the use of micro/mesocosm data and field data to derive benchmarks [10, 12], provided that the resulting data meet certain quality criteria. Unfortunately, there are often insufficient such data of adequate quality to derive benchmark concentrations using only these data. Australia and New Zealand permit such data to be combined with laboratory-based chronic ecotoxicity data to derive GVs [10, 11) using the same method as for laboratory data [11]. As Buchwalter et al. [1] recognise, field data can suffer from confounding and other issues, which make simple cause and effect relationships difficult to elucidate. While they identify some ways of dealing with such issues, it is worth noting that there are other existing and emerging techniques borrowed from other biological disciplines that can be used for determining causality of, and benchmarks for, chemical impacts to single or multiple species measured in the field. A number of these approaches have been described by van Dam et al. [13] and Chariton et al. [14], but more effort is needed to explore their potential and document them as formal tools for field-based WQC derivation.

Buchwalter et al. [1] also recommend that dietary exposures to chemicals be routinely included in toxicity tests and the resulting data used in WQC derivation. They suggest that “it would be good scientific practice to pre-equilibrate the food at various exposure concentrations…thereby more closely reflecting exposures in nature”. Adopting this recommendation will increase the number of potential exposure pathways being considered, but pre-equilibration is not always relevant for highly hydrophilic chemicals or cases where the exposure is as a short pulse. More importantly, this practice could lead to unnaturally high doses of chemicals being ingested due to the very high food loading rates used in many toxicity tests. For example, chronic tests with *Ceriodaphnia* and copepods often use algal cell densities that are orders of magnitude greater than that to which biota would be exposed in the field. Clearly, some thought must be given to how environmentally realistic dietary exposures could be incorporated in benchmark derivation. Nevertheless, we agree that the current inability to include trophic exposure pathways in benchmark derivation means that benchmarks for some chemicals will overestimate the “safe” concentration. This limitation of current toxicity data occurred with the recent derivation of (draft) GVs for perfluorooctane sulfonate (PFOS) and perfluorooctanoic acid (PFOA) in Australia and New Zealand, for which trophic transfer and biomagnification are significant exposure routes. For both of these chemicals, there are very few available data from long-term multigenerational toxicity studies and therefore the GVs will need to be qualified as being potentially under-protective.

Our final thoughts on Buchwalter et al. [1] focus on whether their arguments and associated recommendations would apply equally to legally enforceable WQC/EQS and non-enforceable guidelines? Australia and New Zealand promote an integrated or multiple lines of evidence approach to water quality assessment, where the GVs are not simple pass/fail numbers [10] and have no legal status. Exceedance of a GV indicates there is a moderate to high probability that adverse environmental effects will occur and triggers action – either further site-specific investigation or management action (e.g. reducing the concentration of discharged chemicals). The US criteria and EU EQSs are, more or less, pass/fail numbers and, hence, the quality and confidence in these values are of paramount importance. While the derivation of GVs in Australia and New Zealand still needs to meet minimum quality criteria, the level of confidence in the GVs need not be as high as for WQC/EQS, because the GVs are only one of multiple lines of evidence used to assess water quality. Under these circumstances, it is more beneficial to invest less effort into each default GV so that more can be derived (whilst still meeting the minimum quality standard), but ensuring that specific water quality assessments (e.g., for waste discharges) employ a multiple lines of evidence approach that includes derivation of site-specific GVs where necessary (see examples in [13, 15, 16]). In this way, the effort is allocated at, and tailored to, the relevant spatio-temporal scale of the actual issue, rather than at the high level default GV (national) scale that regardless of the investment in its derivation, is still unlikely to be able to account for important site-specific factors. In line with this, the current revision of Australian and New Zealand Guidelines for Fresh and Marine Water Quality will place an even stronger emphasis on site-specific GVs and multiple lines of evidence approaches to water quality assessment. Thus, while we support the adoption of a multiple lines of evidence approach as outlined by Buchwalter et al. [1], we believe that their arguments and recommendations, while still valid, do not necessarily apply similarly to benchmarks that have no legal basis. This may become an important discussion if international harmonisation of benchmark derivation gains further momentum.

In conclusion, we support the overall approach advocated by Buchwalter et al. [1] and believe that if their approach was implemented it would certainly improve the quality and usefulness of benchmarks. We hope that the additional information provided in our response will enhance their argument while drawing attention to some limitations or points of clarification that will increase its international applicability.

References

[1] Buchwalter DB, Clements WH, Luoma SN. 2017. Modernising water quality criteria in the United States: A need to expand the definition of acceptable data. *Environ Toxicol Chem* 36:285–291.

[2] U.S. EPA. 2016. Invited Expert Meeting on Revising U.S. EPA's Guidelines for Deriving Aquatic Life Criteria. USA: U.S. EPA. [cited 2017 January 30]. Available from: https://www.epa.gov/wqc/invited-expert-meeting-revising-us-epas-guidelines-deriving-aquatic-life-criteria

[3] Merrington G, An YJ, Grist EPM, Jeong SW, Rattikansukha C, Roe S, Schneider U, Sthiannopkao S, Suter GW, van Dam R, van Sprang P, Wang JY, Warne MStJ, Yillia PT, Zhang XW and Leung KMY. 2014. Water quality guidelines for chemicals: learning lessons to deliver meaningful environmental metrics. *Environ Sci Pollut Res* 21:6–16.

[4] Schlekat CE, Merrington G, Leverett D, Peters A. Forthcoming. Chemical standard derivataion for the protection of aquatic life: A guided world tour. *Integr Environ Assess Manag.* DOI: 10.1002/ieam.1919.

[5] Batley GE, Warne MStJ. Forthcoming. Harmonisation of Water and Sediment Quality Guideline Derivation. *Integr Environ Assess Manag*. DOI:

[6] Forbes TL, Forbes VE. 1993. A critique of the use of distribution-based extrapolation models in ecotoxicology. *Funct Ecol* 7: 249–254.

[7] Schudoma D. 1994. Derivation of water quality objectives for hazardous substances to protect aquatic ecosystems: Single-species test approach. *Environ Toxicol Water Qual* 9:263–272.

[8] Warne MStJ. 1998. Critical review of methods to derive water quality guidelines for toxicants and a proposal for a new framework. Supervising Scientist Report 135, Supervising Scientist, Canberra, ACT, Australia.

[9] EC. 2011. Common implementation strategy for the Water Framework Directive (2006/60/EC).

Guidance Document No. 27. Technical guidance for deriving environmental quality standards. European Commission, Brussels, Belgium. p. 41.

[10] ANZECC/ARMCANZ. 2000. National Water Quality Management Strategy, Paper No. 4, Australian and New Zealand Guidelines for Fresh and Marine Water Quality, Volume 1, The Guidelines (Chapters 1–7). Australian and New Zealand Environment and Conservation Council and Agriculture and Resource Management Council of Australia and New Zealand, Canberra, Australia.

[11] Warne MStJ, Batley GE, van Dam RA, Chapman JC, Fox DR, Hickey CW, Stauber JL. 2015. Revised Method for Deriving Australian and New Zealand Water Quality Guideline Values for Toxicants. Updated January, 2017. Prepared for the Council of Australian Government's Standing Council on Environment and Water (SCEW). Department of Science, Information Technology and Innovation, Brisbane, Queensland. [cited 2017 January 15]. Available from https://publications.csiro.au/rpr/pub?pid=csiro:EP159161

[12] Warne MStJ. 2001. Derivation of the ANZECC and ARMCANZ Water Quality Guidelines for Toxicants. *Australas J Ecotoxicol* 7: 123–136.

[13] Van Dam RA, Humphrey CL, Harford AJ, Sinclair A, Jones DR, Davies S, Storey AW. 2014. Site-specific water quality guidelines: 1. Derivation approaches based on physicochemical, ecotoxicological and ecological data. *Environ Sci Pollut Res* 21:118–130.

[14] Chariton AA, Sun M, Gibson J, Webb JA, Leung KMY, Hickey CW, Hose GC. 2015. Emergent technologies and analytical approaches for understanding the effects of multiple stressors in aquatic environments. *Fresh Mar Environ Res* 67: 414–428.

[15] Cormier SM, Paul JF, Spehar RL, Shaw-Allen P, Berry WJ, Suter GW. 2008. Using field data and weight of evidence to develop water quality criteria. *Integr Environ Assess Manag* 4:490–504.

[16] Moore DRJ, Greer CD, Manning G, Woodling K, Beckett KJ, Brain A, Marshall G. Forthcoming. A weight-of-evidence approach to deriving a level of concern for atrazine that is protective of aquatic plant communities. *Integr Environ Assess Manag*. DOI: 10.1002/ieam.1865.